

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-264910

(43)Date of publication of application : 21.09.1992

(51)Int.Cl.

G06F 3/06

G06F 3/06

(21)Application number : 03-026008

(71)Applicant : HITACHI LTD

(22)Date of filing : 20.02.1991

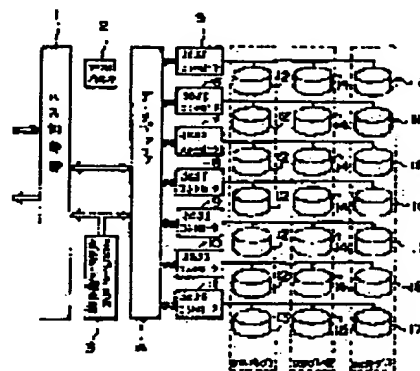
(72)Inventor : NAKANO TOSHIO
MIYAZAKI MICHIO
INOUE YASUO

(54) DATA UPDATING METHOD FOR COMPUTER EXTERNAL STORAGE DEVICE

(57)Abstract:

PURPOSE: To reduce the number of times of access to a storage device at the time of the partial update of data, and to minimize the deterioration of performance by comparing the number of the objective storage devices to be updated and the number of the storage devices other than the object of update with each other, and executing the generation of redundant data, the storage of the update data, and the storage of the updated redundant data by reading out the data from the storage device group of a smaller number.

CONSTITUTION: A bus control part 1 divides a data block sent from a host device in order to store it in plural magnetic disk devices, and a microprocessor 2 controls a whole device. An ECC data generating and data correcting device 3 applies ECC data to the redundant data, and at the time of a data read-out fault, it executes not only the detection of it but the correction of it. Then, the microprocessor 2 calculates the number of objective magnetic disk drives to be updated partially in the ECC group and the number of the magnetic disk drives other than the object to be updated partially in the same ECC group, and operates so as to access the magnetic disk drive of the smaller number.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平4-264910

(43) 公開日 平成4年(1992)9月21日

(51) Int.Cl.⁵

G 0 6 F 3/06

識別記号

3 0 2 E 7165-5B

3 0 4 B 7165-5B

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数1(全5頁)

(21) 出願番号 特願平3-26008

(22) 出願日 平成3年(1991)2月20日

(71) 出願人 00005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 中野 俊夫

神奈川県小田原市国府津2880番地株式会社

日立製作所小田原工場内

(72) 発明者 宮崎 道生

神奈川県小田原市国府津2880番地株式会社

日立製作所小田原工場内

(72) 発明者 井上 靖雄

神奈川県小田原市国府津2880番地株式会社

日立製作所小田原工場内

(74) 代理人 弁理士 秋本 正実

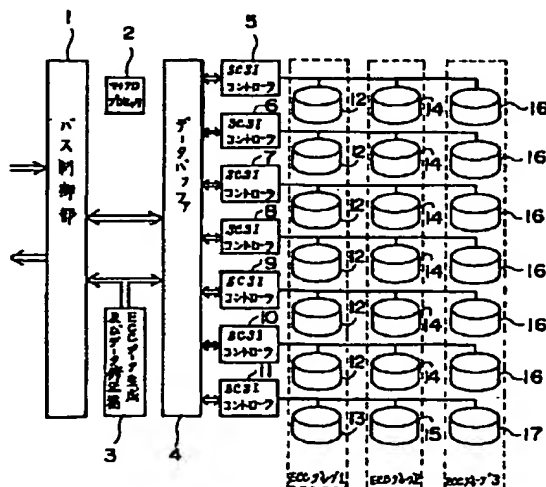
(54) 【発明の名称】 コンピュータ外部記憶装置のデータ更新方法

(57) 【要約】

〔目的〕 データの一部だけを更新する部分更新の際に、記憶装置に対するアクセス回数を低減し、性能劣化を極小化すること。

〔構成〕 部分更新対象磁気ディスクドライブの数と部分更新対象外磁気ディスクドライブの数を比較し、数の少ない方の磁気ディスクドライブからデータを読み出して更新冗長データを作成し、更新データ及び更新冗長データの格納を行う。

【図1】



12~17 磁気ディスクドライブ

【特許請求の範囲】

【請求項1】データブロックを複数のデータに分割し、分割された各データについて冗長データを作成し、上記分割数に等しい数だけ設けられたデータ格納用の記憶装置に分割された各データを格納し、さらに少なくとも1つ設けられた冗長データ格納用の記憶装置に冗長データを格納するコンピュータ外部記憶装置において、上記データ格納用の複数の記憶装置のうちの一部の記憶装置に格納されたデータを更新する際に、更新対象記憶装置の数と更新対象外記憶装置の数を比較し、更新対象記憶装置の数が更新対象外記憶装置の数よりも多いか等しい場合に限って、更新対象であるデータと組合わされて1つのデータブロックを構成するデータを格納している更新対象外記憶装置のアドレスからデータを読み出し、読み出されたデータと更新データとに基づいて更新冗長データを作成し、さらに更新データを上記更新対象記憶装置に書き込み、作成された更新冗長データを上記冗長データ格納用の記憶装置に書き込むことを特徴とするコンピュータ外部記憶装置のデータ更新方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明はコンピュータ外部記憶装置のデータ更新方法にかかり、特にデータを分割して冗長データを付加し、データの分割数及び冗長データ分に等しい複数の記憶装置にデータを格納するデータ格納方法において、複数の記憶装置に記憶されたデータの一部を更新する際に、複数の記憶装置に対するアクセス回数を低減して、性能の低下を極小化するのに好適なコンピュータ外部記憶装置のデータ更新方法に関する。

【0002】

【従来の技術】従来から複数の記憶装置にデータを分割して記憶し、かつ各データについて冗長データを作成して専用の記憶装置に割り当てて記憶する並列データ転送方式が、UCバークレイ校によって提案されている。このような並列データ転送方式は、RAID (Redundant Arrays of Inexpensive Disks) と呼ばれ、高速転送が可能なRAID 3とランダムアクセスに強い多重化／オフラインシークに特徴をもつRAID 5が著名である。以下、磁気ディスク装置を複数台使用したアレイ磁気ディスク装置を例にして、RAID 3とRAID 5について説明する。

【0003】RAID 3は、上位装置のアクセス単位であるデータブロックを複数の磁気ディスクに分割して並列に記録し、データブロック毎に作成される冗長データも合わせて専用の磁気ディスク装置に記録する方式である。そして、上記冗長データは、読み出し障害が発生した際に、データの修復に利用される。RAID 3は、並列にデータ転送するため、単一の磁気ディスク装置と比較して、並列度の分だけ高速転送が可能であり、高性能が期待できる。しかし、その反面、シーク動作が多発す

るランダムアクセスに弱い。また、データの読み出し／更新は、分割してデータを記憶した個々の磁気ディスク装置に対して処理されるのではなく、データブロック単位で行われるのが一般的である。

【0004】RAID 5は、上位装置のアクセス単位であるデータブロックを順次複数の磁気ディスク装置に記録し、あるまとまった単位で冗長データを作成し、この冗長データも合わせて専用の磁気ディスク装置に記録する方式である。RAID 5は、磁気ディスク装置の多重化制御が可能であり、オフラインシークが実現でき、ランダムアクセスに強い反面、データの高速転送は個々の磁気ディスク装置程度というものである。この場合、個々の磁気ディスク装置への個別アクセスが可能である。以上のように、RAID 3、RAID 5はそれぞれ特徴を持っているが、問題点もあり、両者の共存したアレイ磁気ディスク装置も提案されている。

【0005】なお、上記したアレイ磁気ディスク装置の性能向上を目的として、種々の制御方式が提案されている。特に、本発明と関連するものとしては、特開昭62-24481号公報に開示されているように、データを複数の記録媒体に分割して、冗長データ（パリティデータ）を付加して記録し、読み出し時に冗長データにより読み出しエラーを検出し、修正を行うものがある。

【0006】

【発明が解決しようとする課題】一般に、磁気ディスク装置等のコンピュータ外部記憶装置は、ランダムアクセスが可能な低ビットコストの記憶装置という特徴を持つ。前記従来技術において、高速転送という特徴を持つRAID 3の技術は、ランダムアクセスが低性能であるという問題点を持ち、RAID 3及びRAID 5の共存が重要である。RAID 5では、上位装置のアクセス単位であるデータブロックをシリアルに複数の磁気ディスク装置に記憶して行くが、データブロックをさらに大きくとり、RAID 3の技術による高速転送アクセスと、RAID 5による個々の磁気ディスク装置へのランダムアクセスをアクセス特性に応じて切り替えることもできる。ただし、この場合に問題となるのは、データ障害検知または修復用のデータ格納磁気ディスクへのアクセス方法である。

【0007】特開昭62-24481号公報に記載の発明においては、複数の記憶装置を集め、その冗長データを専用の記憶装置に記録し、データ修正のグループを構成している。データ更新の際には、このグループの記憶装置及び冗長データ格納用の記憶装置の当該箇所を読み取り、新冗長データと共に当該グループ内の記憶装置全てにアクセスし、データを更新していた。したがって、RAID 5の技術を持ち込み、当該グループ内の一部分の記憶装置を部分更新するときには、まず更新対象の記憶装置と冗長データ格納用の記憶装置の当該箇所を読み取り、新冗長データと共に当該記憶装置に対してデータ

を更新する必要があった。

【0008】本発明の目的は、データの一部分だけを更新する部分更新の際に、記憶装置に対するアクセス回数を低減し、性能劣化を極小化しつつ、高速データ転送を実現する事が可能なコンピュータ外部記憶装置のデータ格納方法を提供することにある。

【0009】

【課題を解決するための手段】本発明のコンピュータ外部記憶装置のデータ更新方法は、データブロックを複数のデータに分割し、分割された各データについて冗長データを作成し、上記分割数に等しい数だけ設けられたデータ格納用の記憶装置に分割された各データを格納し、さらに少なくとも1つ設けられた冗長データ格納用の記憶装置に冗長データを格納するコンピュータ外部記憶装置に適用されるものであり、次の特徴を有する。

【0010】すなわち、上記データ格納用の複数の記憶装置のうちの一部の記憶装置に格納されたデータを更新する際に、まず更新対象記憶装置の数と更新対象外記憶装置の数を比較する。そして、更新対象記憶装置の数が更新対象外記憶装置の数よりも多いか等しい場合に限って、更新対象であるデータと組合わされて1つのデータブロックを構成するデータを格納している更新対象外記憶装置のアドレスからデータを読み出す。次に、読み出されたデータと更新データとに基づいて更新冗長データを作成し、さらに更新データを上記更新対象記憶装置に書き込み、作成された更新冗長データを上記冗長データ格納用の記憶装置に書き込む。これによって、データの部分更新が行われる。

【0011】

【作用】本発明によれば、複数の記憶装置のうちの一部の記憶装置においてデータ更新する際に、更新対象の記憶装置の数と更新対象外の記憶装置の数を比較して、数が少ない方の記憶装置群からデータを読み出して、更新冗長データの作成、更新データの格納、更新冗長データの格納を行うため、データの部分更新の際に、記憶装置に対するアクセス回数を低減し、性能劣化を極小化し、これによって高速データ転送を実現する事ができる。

【0012】

【実施例】以下、添付の図面に示す実施例により、更に詳細に本発明について説明する。図1は本発明をアレイド磁気ディスク装置に適用した場合の一実施例を示すブロック図である。

【0013】図1において、1は上位装置から転送されてくるデータブロックを複数の磁気ディスク装置に格納するために分割するバス制御部、2は装置全体を制御するマイクロプロセッサ、3は本例では冗長データにECCデータを適用し、データ読み出し障害時、その検知だけではなく修正まで行うECCデータの生成及びデータ修正器、4は分割したデータを一時的に貯え同期合わせをするためのデータバッファ、5から10はバス制御部

1において複数に分割されたデータブロックを各磁気ディスクドライブ12、14、16に書き込み、又はその逆に各磁気ディスクドライブ12、14、16からのデータの読み出しを制御するSCSIコントローラ、11はECCデータ生成及びデータ修正器3で作成されたECCデータを書き込み、又はその逆にECCデータの読み出しを制御するSCSIコントローラ、12は第1の分割されたデータブロックを記憶する磁気ディスクドライブ、13は第1の分割されたデータブロックのECCデータを格納する磁気ディスクドライブ、14は第2の分割されたデータブロックを格納する磁気ディスクドライブ、15は第2の分割されたデータブロックのECCデータを格納する磁気ディスクドライブ、16は第3の分割されたデータブロックを記憶する磁気ディスクドライブ、17は第3の分割されたデータブロックのECCデータを格納する磁気ディスクドライブである。ここで、第1～第3とは、図1中に破線で示すように、データ修復の単位である複数の磁気ディスクドライブから構成されるECCグループ1～3に対応している。

【0014】図示しない上位装置は、図1に示すアレイド磁気ディスク装置を1つ以上の複数の論理ディスクボリュームとして認識する。そして、そのアクセス領域は、本例においては、ECCグループ1、2、3から構成される3個の磁気ディスクドライブ群によって分割されている。論理ディスクボリュームと磁気ディスクドライブとの対応は、マイクロプロセッサ2の指定によってバス制御部1が行う。

【0015】次に、一般的なデータの書き込み/読み出しについて述べる。アレイド磁気ディスク装置の制御方式としては、UCパーレイ校提案のRAID3、RAID5という2種類の制御方式が著名である。本実施例のアレイド磁気ディスク装置では両方式の使い分けを行い、高速転送が可能なRAID3、及びランダムアクセスに強いRAID5の両者の特徴を兼ね備える。両制御方式の使い分けは、上位装置からのコマンド指示による。

【0016】RAID3方式においては、ECCグループ単位にアクセスされ、ECCグループはデータ修復の単位である複数の磁気ディスクドライブから構成されている。このRAID3方式においてデータ更新を行う場合、バス制御部1がデータブロックを受領すると、バス制御部1は当該データブロックを複数の固定長ブロックに分割する。このとき、ECC生成及びデータ修正器3はECCデータを生成し、データ読み出しの際の信頼性を向上させる。これら細分化されたデータは各磁気ディスクドライブに格納される。一方、読み出しの際も、同様にECCグループ単位で磁気ディスクドライブにアクセスし、データを読み出す。

【0017】個々の磁気ディスクドライブに複数バイトづつデータを書き込む場合には、次のように動作する。

すなわち、本例では、上位装置が3072(512パイ

ト×6=3072)バイト単位でデータを書き込む場合、上位装置は3027バイト単位でアクセスし、個々の磁気ディスクドライブは512バイト単位でデータを格納する。このように、個々の磁気ディスクドライブをアクセスするケースも考えられる。

【0018】また、本例では、上位装置が512バイト単位でもアクセスする場合も考えられる。この場合には、上位装置は、磁気ディスクドライブの多重制御が可能なRAID5を指定する。つまり、データ修復のため、複数の磁気ディスクドライブから構成される同一ECCグループの中の一部の磁気ディスクドライブのみを更新する部分更新の場合である。

【0019】次に、図3に示す従来方式について説明し、図2に示す本発明の方法の例と比較する。図3に示す従来例では、斜線で示す4個の磁気ディスクドライブが更新対象であり、同一ECCグループ中のECCドライブと合わせ、対象箇所のデータ読み出しが行われる。この読み出しを図中Aで示す。これにより、磁気ディスクドライブからのデータの読み出しを5回行うことになる。次に、読み出されたデータと更新データに基づいて、更新ECCデータの作成が行われる。この工程を図中Bで示す。最後に4個の磁気ディスクドライブの当該アドレスに更新データを、ECC用磁気ディスクドライブに更新ECCデータを書き込む。この書き込みを図中Cで示す。これにより、磁気ディスクドライブへのデータ書き込みを5回行うことになる。以上に説明したように、従来方式では、本例の場合、10回の磁気ディスクドライブへのアクセスが必要であった。

【0020】同様に、図2を用いて、本発明のデータ格納方式を説明する。本発明の方式でも、図3と同様に、4個の磁気ディスクドライブが更新対象である。しかし、図1に示すマイクロプロセッサ2は、ECCグループ内の部分更新対象磁気ディスクドライブの数と、同一ECCグループ内の部分更新対象外磁気ディスクドライブの数を算出する。そして、数の少ない方をアクセスするように動作する。この場合、マイクロプロセッサ2は、部分更新対象磁気ディスクドライブをアクセスしてデータ更新した場合のアクセス回数と、部分更新対象外磁気ディスクドライブをアクセスしてデータ更新した場合のアクセス回数を比較し、アクセス回数の少ないほうを選択するようにしても良い。

【0021】本例の場合、部分更新対象ディスクドライブは4個であるのに対し、部分更新対象外磁気ディスクドライブは2個であり、アクセス回数の少ない部分更新対象外磁気ディスクドライブのアクセスが選択される。したがって、マイクロプロセッサ2は、部分更新対象外の2個の磁気ディスクドライブを読み出す指示を出し、

データバッファ4に当該アドレスのデータが読み出される。ここで、部分更新対象外の2個の磁気ディスクドライブから読み出されるデータは、更新対象であるデータと組合わされて1つのデータブロックを構成しているデータを格納しているアドレスに記憶されているデータである。この読み出しを図中Dで示す。これにより、磁気ディスクドライブからのデータの読み出しを2回行うことになる。次に、読み出されたデータと更新データとに基づいて、新更新ECCデータを生成する。この工程を図中Eで示す。最後に、読み出していない4個の磁気ディスクドライブ(部分更新対象磁気ディスクドライブ)の当該アドレスに更新データを書き込み、ECC用の磁気ディスクドライブに更新ECCデータを書き込む。この書き込みを図中Fで示す。これにより、磁気ディスクドライブへのデータ書き込みを5回行うことになる。

【0022】以上の説明から明らかなように、本実施例によれば、磁気ディスクドライブへのアクセス回数が7回であり、従来例と比較しての磁気ディスクドライブへのアクセスが3回分低減できたことになる。

【0023】本実施例によれば、データ修復のため、複数の記憶装置から構成されるグループの中の一部の記憶装置のみを更新する部分更新の際に、磁気ディスクドライブへのアクセス回数を低減できるので、性能向上に効果があり、かつECCデータを確実に書き込むことができる。

【0024】

【発明の効果】本発明によれば、データの一部だけを更新する部分更新の際に、記憶装置に対するアクセス回数を低減し、性能劣化を極小化しつつ、高速データ転送を実現する事が可能になる効果がある。

【図面の簡単な説明】

【図1】本発明をアレイ磁気ディスク装置に適用した場合の一実施例を示すブロック図。

【図2】本発明の方法を用いて複数の磁気ディスクドライブの内の一部分の磁気ディスクドライブのみを更新する部分更新を示す説明図。

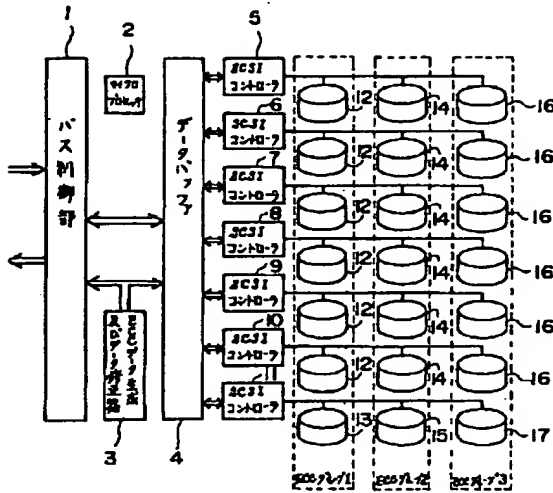
【図3】従来の方法を用いて複数の磁気ディスクドライブの内の一部分の磁気ディスクドライブのみを更新する部分更新を示す説明図。

【符号の説明】

- 1 バス制御部
- 2 マイクロプロセッサ
- 3 ECCデータの生成及びデータ修正部
- 4 データバッファ
- 5~11 SCSIコントローラ
- 12~17 磁気ディスクドライブ

【図1】

【図1】



12~17 ----- 磁気ディスクドライブ

【図3】

【図3】

